Efficiency of cleaning and disinfection of surfaces: correlation between assessment methods

ABSTRACT

Objective: to assess the correlation among the ATP-bioluminescence assay, visual inspection and microbiological culture in monitoring the efficiency of cleaning and disinfection (C&D) of high-touch clinical surfaces (HTCS) in a walk-in emergency care unit. Method: a prospective and comparative study was carried out from March to June 2015, in which five HTCS were sampled before and after C&D by means of the three methods. The HTCS were considered dirty when dust, waste, humidity and stains were detected in visual inspection; when ≥2.5 colony forming units per cm² were found in culture; when ≥5 relative light units per cm² were found at the ATP-bioluminescence assay. Results: 720 analyses were performed, 240 per method. The overall rates of clean surfaces per visual inspection, culture and ATP-bioluminescence assay were 8.3%, 20.8% and 44.2% before C&D, and 92.5%, 50% and 84.2% after C&D, respectively (p<0.001). There were only occasional statistically significant relationships between methods. Conclusion: the methods did not present a good correlation, neither quantitative nor qualitatively.

Descriptors: Equipment Contamination; Housekeeping; Nursing Audit; Health Facility Environment; Infectious Disease Transmission.
INTRODUCCIÓN

Aunque existen esfuerzos crecientes para prevenir y elevar la conciencia sobre la atención posoperatoria, esta sigue siendo una de las mayores amenazas a la práctica clínica. Las áreas de atención clínica altamente tocadas (HTCS) pueden ser fuentes de transmisión de microorganismos y enfermedades infecciosas, lo que puede aumentar el riesgo de infecciones nosocomiales (IN). En este sentido, se hace necesaria la aplicación de prácticas de limpieza y desinfección (L&D) adecuadas para prevenir la transmisión de microorganismos y enfermedades infecciosas. Sin embargo, la falta de estudios sobre la correlación entre los métodos de L&D y la presencia de microorganismos en superficies clínicas altamente tocadas (SCAT) en un entorno de atención clínica es un desafío a la hora de seleccionar el método más adecuado para medir la eficacia de las prácticas de L&D. En este sentido, el objetivo del estudio fue valorar la correlación entre la presencia de ATP en superficies de HTCS y la visión visual y el cultivo microbiológico en un entorno de atención clínica.

MÉTODOS

El estudio fue un estudio comparativo, prospectivo, realizado de marzo a junio de 2015, cuando cinco SCAT fueron muestreadas antes y después de L&D de rutina por los tres métodos. Las SCAT fueron consideradas sucias cuando presentaban: visión visual: polvo, derrames, humedad y manchas; en cultivo: ≥ 205 unidades formadoras de colonias por cm², y en ATP-bioluminiscencia: ≥ 5 Unidades Relativas de Luz por cm². Los SCAT fueron realizadas 720 evaluaciones, 240 por método. La tasa global de superficies limpias por visión visual, cultivo y ATP-bioluminiscencia fue, respectivamente, 8,3%, 20,8% y 44,2% antes de la L&D y de 92,5%, 50% y 84,2% después (p<0,001). Existen sólo asociaciones puntuales estadísticamente significativas entre los métodos. Conclusion: los métodos no presentan buena correlación cuantitativa, ni cualitativa.

RESULTADOS

Se realizaron 720 evaluaciones, 240 por método. La tasa de superficies limpias por visión visual, cultivo y ATP-bioluminiscencia fue, respectivamente, 8,3%, 20,8% y 44,2% antes de la L&D y de 92,5%, 50% y 84,2% después (p<0,001). Existenten sólo asociaciones puntuales estadísticamente significativas entre los métodos. Conclusion: los métodos no presentan buena correlación cuantitativa, ni cualitativa.

DESCRIPTORES: Contaminación de Equipos; Servicio de Limpieza; Auditoría de Enfermería; Ambiente de Instituciones de Salud; Trasmisión de Enfermedad Infecciosa.
Routine of cleaning and disinfection of surfaces

The C&D of the surfaces surveyed was performed by the nursing team once a day – at the beginning of the morning shift – or when organic liquids were spilled over, except for the mattress, which was always disinfected after patient discharge. The procedure included the use of cleaning cloths folded into four parts, sprinklers and a hospital disinfectant for fixed surfaces, composed of glucoprotamin (12.4%) and alkyl dimethyl benzyl ammonium chloride (15%), which acts as a detergent and disinfectant, therefore cleaning and disinfecting at one go.

Data collection methods and cleaning standards

The HTCS were sampled by visual inspection, microbiological culture and ATP-bioluminescence assay before and after C&D. The surfaces were sampled – exclusively by the authors of this study, and only once a day – right before and 10 minutes after the end of the C&D morning session\(^4\). There was not a specific department for collection of HTCS.

Visual inspection

First method applied: surfaces that contained dust, waste (blood, wound ooze, organic liquids, physiological serum crystals, ointment or cream, oil, solute, etc.), patch residues, humidity and stains were considered dirty\(^4\).\(^5\).

ATP-bioluminescence assay

In order to detect ATP by bioluminescence, a hand-held luminometer was used (Clean-Trace ATP System; 3M\(^4\)) and a Clean Trace Kit - a specific swab. Following the manufacturer’s recommendation and the literature\(^6\),\(^7\),\(^8\),\(^9\), a sterile cotton swab was leaned over the tested surface until it was slightly bend-ed, forming a 30° angle, and then rubbed in zigzag fashion. This method measures the amount of organic ATP found in the sample. By means of a proper swab, the organic material found on the surface is collected and transferred to a detection device made up of an enzyme-substrate compound (luciferin-luciferase). The reaction that results from the contact of the sample with this compound releases a certain type of light, whose intensity is measured by hand-held luminometers and which is expressed in RLU. The amount of RLU is proportional to the amount of ATP, which in turn is proportional to the density of the organic material\(^6\)-\(^8\).\(^9\). The surfaces were considered clean when the ATP index was < 5 RLU/cm\(^2\), collected from a 100 cm\(^2\) surface, that is, < 500 RLU/surface\(^3\)-\(^5\),\(^9\).

Microbiological cultures

The microbiological samples were collected by means of RODAC plates (Replicate Organism Detection And Counting) with tryptcase soy agar (TSA), which contained sanitizer neutralizers and a 24 cm\(^2\) area\(^3\),\(^10\). The plates were labeled, pressed for 10 seconds against the surfaces at ~ 25 g/cm\(^2\), without any sideways movement, and incubated at 37°C for 24-48h\(^6\),\(^10\)-\(^11\). For the aerobic colony counting (ACC), an electronic and digital colony counter was used (Logen\(^\) LS60000). The surfaces were considered clean when the ACC was < 2.5 CFU/cm\(^2\), that is, < 60 CFU/plate\(^3\)-\(^5\),\(^10\).

Analysis of results and statistics

The comparison between the methods and associations between the disapproval rates was analyzed by means of a Fisher’s exact test or a Kruskall-Wallis test. The levels of RLU and ACC obtained before and after C&D were analyzed by means of the Wilcoxon signed-rank test. The Spearman’s correlation test was used to examine the correlation between the ATP and ACC scores, and the Mc Nemar’s test was used for the qualitative correlation (approved or disapproved). The receiver operating characteristic (ROC) curve was drawn, and the significance level adopted was 5% (p<0.05).

RESULTS

A total of 720 samples were collected, 240 of which by monitoring, half of them before and the other half after C&D. Each one of the five surfaces was sampled 48 times per method, 24 of which were done before C&D and the other 24 after C&D. According to Table 1, out of the 120 evaluations carried out before C&D, 8.3%, 20.8% and 44.2% were considered clean, respectively, by visual inspection, ACC, and ATP-bioluminescence, against 92.5%, 50% and 84.2% after C&D (p<0.001).

Considering that the p-value is < 0.05, which shows a statistically significant difference, and that we analyzed the correlation among methods, p ≥ 0.05 indicates the existence of a correlation among methods, and p = 1.00, a perfect correlation. Thus, when the disapproval rates of all surfaces are compared (Table 1), we can see that there is only one significant association (p>0.05): ACC and visual inspection after C&D. In all, there were 16 associations (six between ATP and visual inspection; five between ATP and ACC, and five between ACC and visual inspection) and 20 discrepancies. Regarding the cut-off points, the ATP-bioluminescence assay showed scores that were higher than those of ACC, both before and after cleaning and/or disinfection.

With the Spearman’s correlation test (data not presented), there was no statistically significant correlation between ACC and RLU, which indicates that when there is a decrease in CFU after cleaning, it is not possible to assume a decrease in RLU, and vice-versa. There was a great variation between the correlation coefficients: from -0.611 to 0.905. Additionally, the Mc Nemar’s test was used to assess the qualitative discrepancy (dirty or clean) among methods. Qualitative results obtained before and after C&D were considered. The test results were highly significant, suggesting that there is a discrepancy among the methods with regard to approval or disapproval of surfaces: ATP-bioluminescence vs. ACC (p<0.0001); ATP-bioluminescence vs. visual inspection (p<0.0001); and ACC vs. visual inspection (p=0.0006).

According to Figure 1, in general, the levels of ATP and ACC were significantly lower in post-C&D than in pre-C&D. The medication preparation area 2 was the only HTCS for which there was not a significant decrease in the ATP counting after C&D. The gray asterisks represent outliers, that is, observation points that are far from the other scores. The presence of many outliers may have hampered the interpretation...
of results of the statistical tests applied. However, we can see that the median and the 25-75 interquartile range were considerably lower after C&D for all surfaces.

Adopting the reference ACC < 2.5 CFU/cm² for the definition of a clean surface[3,9-10], following the ROC curve, the best cut-off point for ATP-bioluminescence was 7.9 RLU/cm². The characteristics of visual inspection and ATP-bioluminescence with a < 5 RLU/cm² cut-off point and < 8 RLU/cm² with regard to the microbiological comparison are shown in Table 2.

### Table 1 – Disapproval rates, median and variation, according to collection time, surface and monitoring method

<table>
<thead>
<tr>
<th>Time / surfaces</th>
<th>Visual Disapproval n (%)</th>
<th>ATP*(RLU/cm²) Median (variation)</th>
<th>ACC (CFU/cm²) Disapproval n (%)</th>
<th>p value*</th>
<th>ATP vs visual</th>
<th>ATP vs ACC</th>
<th>ACC vs visual</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before C&amp;D</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med Prep 1</td>
<td>23 (95.8)</td>
<td>6.9 (1.2-209.7)</td>
<td>15 (62.5)</td>
<td>3.1 (0.1-74.6)</td>
<td>13 (54.2)</td>
<td>0.010</td>
<td>0.770†</td>
</tr>
<tr>
<td>Heart Monitor</td>
<td>16 (66.7)</td>
<td>8.1 (2-1148.7)</td>
<td>15 (62.5)</td>
<td>1.4 (0.9-11)</td>
<td>8 (33.3)</td>
<td>1.000‡</td>
<td>0.082‡</td>
</tr>
<tr>
<td>Med Prep 2</td>
<td>23 (95.8)</td>
<td>11.2 (3.3-37.5)</td>
<td>22 (91.7)</td>
<td>2.9 (0.1-67.7)</td>
<td>13 (54.2)</td>
<td>1.000‡</td>
<td>0.008</td>
</tr>
<tr>
<td>Dressing trolley</td>
<td>24 (100)</td>
<td>14 (4.1-2920.3)</td>
<td>23 (95.8)</td>
<td>3.5 (0.4-69)</td>
<td>16 (66.7)</td>
<td>1.000‡</td>
<td>0.023</td>
</tr>
<tr>
<td>Mattress</td>
<td>24 (100)</td>
<td>15.3 (1.9-221.2)</td>
<td>20 (83.3)</td>
<td>2.9 (0.5-82)</td>
<td>13 (54.2)</td>
<td>0.109‡</td>
<td>0.060‡</td>
</tr>
<tr>
<td>All</td>
<td>110 (91.7)</td>
<td>11.2 (1.2-2920.3)</td>
<td>95 (87.1)</td>
<td>2.7 (0.1-81.9)</td>
<td>67 (61.4)</td>
<td>0.010</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>After C&amp;D</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med Prep 1</td>
<td>1 (4.2)</td>
<td>4.2 (0.4-18.5)</td>
<td>10 (41.7)</td>
<td>0.9 (0-18.8)</td>
<td>2 (8.3)</td>
<td>0.004</td>
<td>0.017</td>
</tr>
<tr>
<td>Heart Monitor</td>
<td>0 (0)</td>
<td>1.4 (0.5-3.7)</td>
<td>0 (0)</td>
<td>0.2 (0-10.5)</td>
<td>3 (12.5)</td>
<td>1.000‡</td>
<td>0.234‡</td>
</tr>
<tr>
<td>Med Prep 2</td>
<td>1 (4.2)</td>
<td>7.5 (0.7-56.7)</td>
<td>16 (66.7)</td>
<td>0.4 (0.1-8)</td>
<td>3 (12.5)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dressing trolley</td>
<td>9 (37.5)</td>
<td>7.1 (0.5-21.5)</td>
<td>18 (75)</td>
<td>0.1 (0-68.5)</td>
<td>4 (16.7)</td>
<td>0.534‡</td>
<td>0.724‡</td>
</tr>
<tr>
<td>Mattress</td>
<td>1 (4.2)</td>
<td>5.9 (1.4-271)</td>
<td>14 (58.3)</td>
<td>0.8 (0-5.4)</td>
<td>2 (8.3)</td>
<td>&lt;0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>All</td>
<td>12 (10.0)</td>
<td>4.9 (04-270.9)</td>
<td>58 (48.3)</td>
<td>0.2 (0-68.5)</td>
<td>14 (11.7)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Notes: *ATP-bioluminescence; †Test for two proportions; ‡p ≥0.05, indicating a correlation between methods; ATP – adenosine triphosphate; RLU – relative light units; ACC – aerobic colony count; CFU – colony forming units; C&D – cleaning and disinfection.

### Table 2 – Characteristics of non-microbiological methods, according to the comparison with a microbiological culture for the definition of a clean or dirty surface

<table>
<thead>
<tr>
<th>Method</th>
<th>Characteristics of the test to define a surface as dirty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity</td>
</tr>
<tr>
<td>Visual inspection</td>
<td>83.1</td>
</tr>
<tr>
<td>ATP*(&lt; 5 RLU/cm²)</td>
<td>78</td>
</tr>
<tr>
<td>ATP*(&lt; 8 RLU/cm²)</td>
<td>62.3</td>
</tr>
</tbody>
</table>

Notes: *ATP-bioluminescence; †positive predictive value; ‡negative predictive value.
DISCUSSION

This is the first research study carried out by the authors comparing methods for effectively monitoring C&D of surfaces in out-of-hospital environments. Generally speaking, the results showed a weak association between qualitative results (clean or dirty) of ATP-bioluminescence and the other monitoring methods. Both visual inspection and ATP-bioluminescence did not show any correlation with microbiological comparison when they were cross-checked in different ways and by statistical tests. In all methods, the quantitative and/or qualitative results indicated that routine C&D had positive effects on sanitation of the surfaces surveyed.

According to the ROC curve, the best cut-off point of ATP-bioluminescence for the referred unit was 8 RLU/cm². With that cut-off point, there is an increase in sensitivity, a decrease in specificity and an increase in accuracy (Table 2).

Although the cleaning of surfaces is internationally considered as necessary to control HAI, so far there is no consensus that is universally accepted regarding the preferred methods to assess C&D of hospitals, much less so in an WECU.(12). The efficiency of ATP-bioluminescence has been tested in many studies(3,4,12-13), and strong correlations between ACC and levels of ATP were found under specific and controlled conditions, but they were limited in healthcare environments(1,3,14). This can be explained by the diversity of ATP measurement systems, cut-off points and surface C&D practices among the institutions/units where those studies were conducted. In addition, if cleaning is not appropriate, dirt and microorganisms cannot be removed, but only relocated, which leads to the discrepancy with microbiological cultures(3).

Another explanation can be ATP stability. A study proved that, in the absence of C&D, ATP residues coming from both organic material and microorganisms (dead or alive), do not deteriorate rapidly. After 29 days, surfaces infected with suspension of *P. aeruginosa*, *E. faecalis* and *C. albicans* kept ATP levels of 65%, 69% and 96% of levels originally present in the solution, respectively. Surfaces that contained blood had 100% and 8% of the original ATP after 4 and 29 days, respectively(15). Therefore, when C&D methods fail to remove dirt, ATP can remain stable for more than 24 hours on ambient surfaces; and microorganisms are then dead by the action of the disinfectant. Thus, after C&D, there will be a high RLU score and a low ACC.

Although technologies that identify microbial and non-microbial ATP are commercially available, it is important to highlight that most studies carried out in health units use a bioluminescence system that measures total ATP (organic material, dead or living microorganisms), whereas microbiological cultures measure viable microorganisms. A surface may contain organic material in abundance, but not necessarily...
a high microbial density and vice-versa. In that sense, the interpretation of ATP can never be made so as to indicate the presence or absence of pathogens responsible for HAI\(^9\).

Previous studies\(^16-17\) recommended to find corresponding benchmarks of cleanliness in a certain configuration, and use ATP-bioluminescence to monitor the ability of an intervention in reducing the amount of microbes to a predefined level. In other words, ATP-bioluminescence is not a reliable method for the identification and surveillance of potentially infected HTCS, but rather a tool that assesses the efficiency of cleaning procedures or other infection controlling measures.

In addition to the discrepancy with the microbiological comparison, a recent study\(^18\) reported the absence of a correlation between two different measurement systems of ATP-bioluminescence (Kikkoman ATP device with Luchipak-Pen swabs and Hygiene ATP device with Ultrasnap swabs). In addition to the intrinsic singularities of each system, the distribution of organic material over the surface (some areas may have more dirt on a same surface) and the researcher’s ability to keep the homogeneity during the collection with the swab may interfere in RLU interpretation.

Many studies have found a weak correlation between non-microbiological methods and ACC for the monitoring of the efficiency of C&D in HTCS\(^4,13,16-19\). In this study, visual inspection and ATP-bioluminescence differed from ACC qualitative results in 29.6\% and 45.8\% of cases, respectively. Similar results were found in a study\(^19\) that used a hand-held luminometer (CleanTrace ATP System; 3M\(^\) ) and a Clean Trace kit and cut-off point of <5 RLU/cm\(^2\) to assess the efficiency of final cleaning, in which the discrepancy of visual inspection and ATP-bioluminescence with the microbiological comparison was 42\% and 37\%, respectively. These results support the fact that non-microbiological methods cannot estimate the efficiency of cleaning when microbiological cultures are used as a reference.

Although visual inspection has a weak correlation with ACC, as documented in many studies\(^3-5,9\), this study and in another one it is the method that is closest to microbiological comparison, as far as sensitivity, specificity and accuracy values are concerned. This can be related, to some extent, to the evaluator’s skills. The greatest limiting factor is its low specificity, which varies from 9\% to 65\%\(^13,19\). In addition, it is known that, after C&D procedures, many more surfaces considered as dirty by ACC are seen as clean by visual inspection when compared to another non-microbiological method (ATP-bioluminescence, fluorescent marker)\(^4,13\).

Studies suggest that visual assessment is not sufficient to ensure the quality of the process, and it is less efficient than quantitative methods in raising awareness about the need to improve sanitation practices\(^9\). The use of ATP-bioluminescence to assess and have feedback of results leads to greater awareness, collaboration, communication and education of cleaning and nursing staffs, and therefore, it improves the efficiency of practices\(^21-22\).

In this study, sensitivity, specificity, accuracy, PPV and NPV of visual inspection and ATP-bioluminescence were calculated using a microbiological culture (cut-off point of 2.5 CFU/cm\(^2\)) as a “reference” test. The results obtained were very similar to those of a study conducted in the wards of a teaching hospital\(^13\), confirming that sensitivity and specificity of ATP-bioluminescence are around 80\% and 40\%, respectively. The referred study found a very low specificity for visual inspection (9\%), suggesting that infected surfaces can be approved by this method, especially after C&D.

However, it should be highlighted that these findings are far from being unanimous in the literature. Another study\(^20\) found different results, in which sensitivity of visual inspection was 27.3\% and specificity, 94.6\%. This might be related to the fact that there is no standardization of techniques, supplies, cut-off points and technology of the monitoring methods.

It was observed that ATP-bioluminescence <8 RLU/cm\(^2\), the cut-off point suggested by the ROC curve, reduced sensitivity but increased specificity, PPV, and therefore, the accuracy of the test, when compared to the cut-off point <5 URL/cm\(^2\) that is conventionally used for cleaning of hospital environments\(^4,5,9\).

**Study limitations**

The limitations of this study concern its development in only one institution, the small sample considered and the study design, which did not allow for establishing a relationship between the results of the three methods before and/or after C&D with the contraction of HAI. Although the surfaces sampled by swab (ATP-bioluminescence assay) and RODAC plates before and after C&D were adjacent, it is possible that different levels of dirt may have been present in different areas of the surface. The sample was defined for convenience, but considering that we chose the rooms with greater risks of HAI and clinical surfaces that are more frequently touched by hands and related to invasive procedures, we believe that the effect of the non-randomized sampling of rooms and surfaces on the internal validity and/or mainstreaming is minor. From the perspective of cross-transmission of pathogens, this was the best sample possible for this study. Finally, although there are limited data about the cut-off points that are clinically relevant to reduce the transmission of pathogens, cut-off points for each test were used on the basis of studies carried out in hospitals, which may not be suitable to a WECU.

**Contributions to the nursing, healthcare or public policy fields**

The results of this study provide inputs for evidence-based nursing practice, since they provide grounds for the choice of the ideal method of assessment of the efficiency of C&D of HTCS in daily situations, outbreaks and assessment of the impact of specific interventions, as well as they indicate the best cut-off points (benchmarks) of ATP-bioluminescence assay for a WECU, according to its singularities and using a specific and widely known device to measure the bioluminescence. In addition, this study fosters: the creation of public health policies focused on patient safety, with regard to the ongoing or implemented assessments of C&D of surfaces, as well as operational guidelines; health education actions (continuing education, curricular structuring of undergraduate and graduate courses); and future research to address other inputs, ways and devices for monitoring C&D.

**CONCLUSIONS**

No consistent correlation was found among ATP-bioluminescence, visual inspection and ACC. Thus, visual inspection...
performed in a systematic and standardized way can be the preferable method to assess the efficiency of C&D routine in a WECU, given its sensitivity and accuracy, as well as its simplicity, ease, low cost and minimal training required. However, when it is intended to improve C&D practices, audit C&D efficiency, provide immediate feedback on results and assess the impact of specific interventions – educational, operational, logistic –, the use of an ATP-bioluminescence assay and/or microbiological cultures is recommended.

**FUNDING**

Foundation for the Support of Teaching, Science and Technology Development of the state of Mato Grosso do Sul (FUNDECT) for the Unified Health System (SUS) - FUNDECT/DECIT-MS/CNPq/SES call, agreement 04/2013 PPSUS-MS, protocol n. 26434.386.4552.26042013, (FUNDECT; funding of research) and Support and Evaluation of Graduate Education (CAPES - PhD scholarship from first author).

### REFERENCES


